



Apache Kylin在百度地图的实践

百度地图开放平台业务部 - 王冬

目录

大数据多维分析平台建设背景

大数据查询框架选型及调研

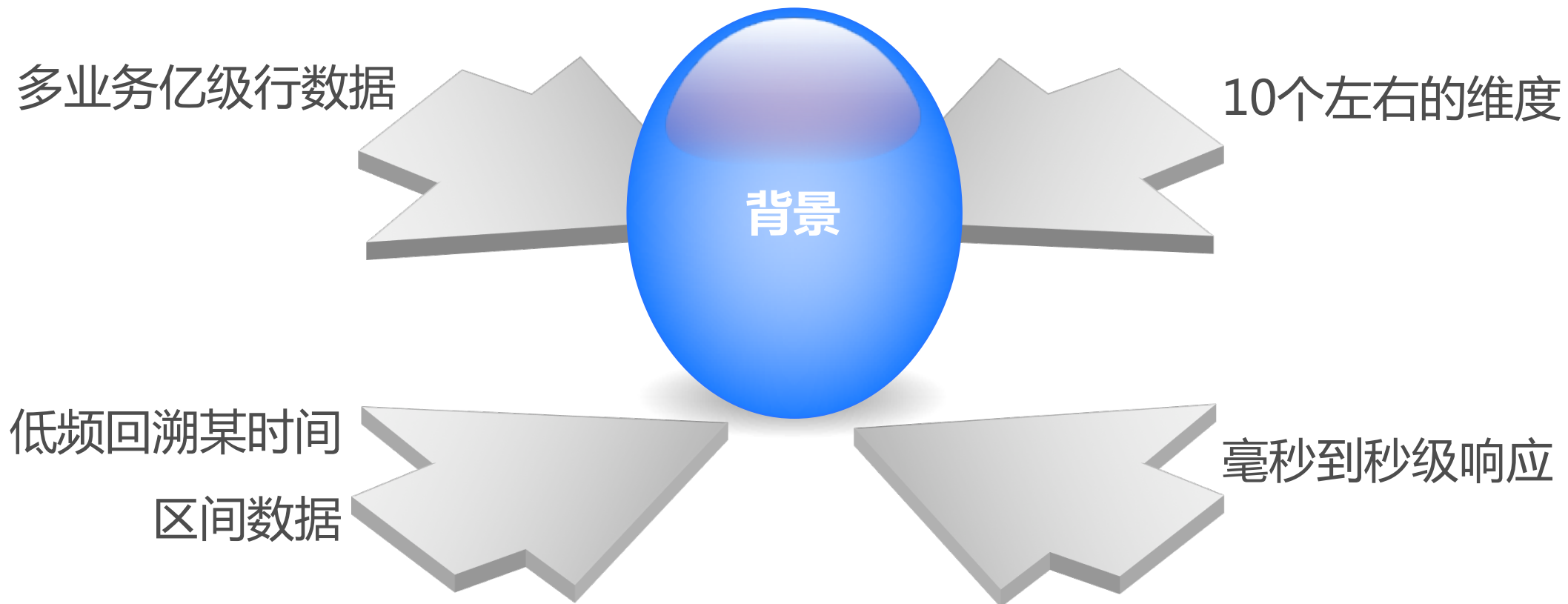
大数据OLAP多维分析平台建设

我们的深度开发

我们的项目实践

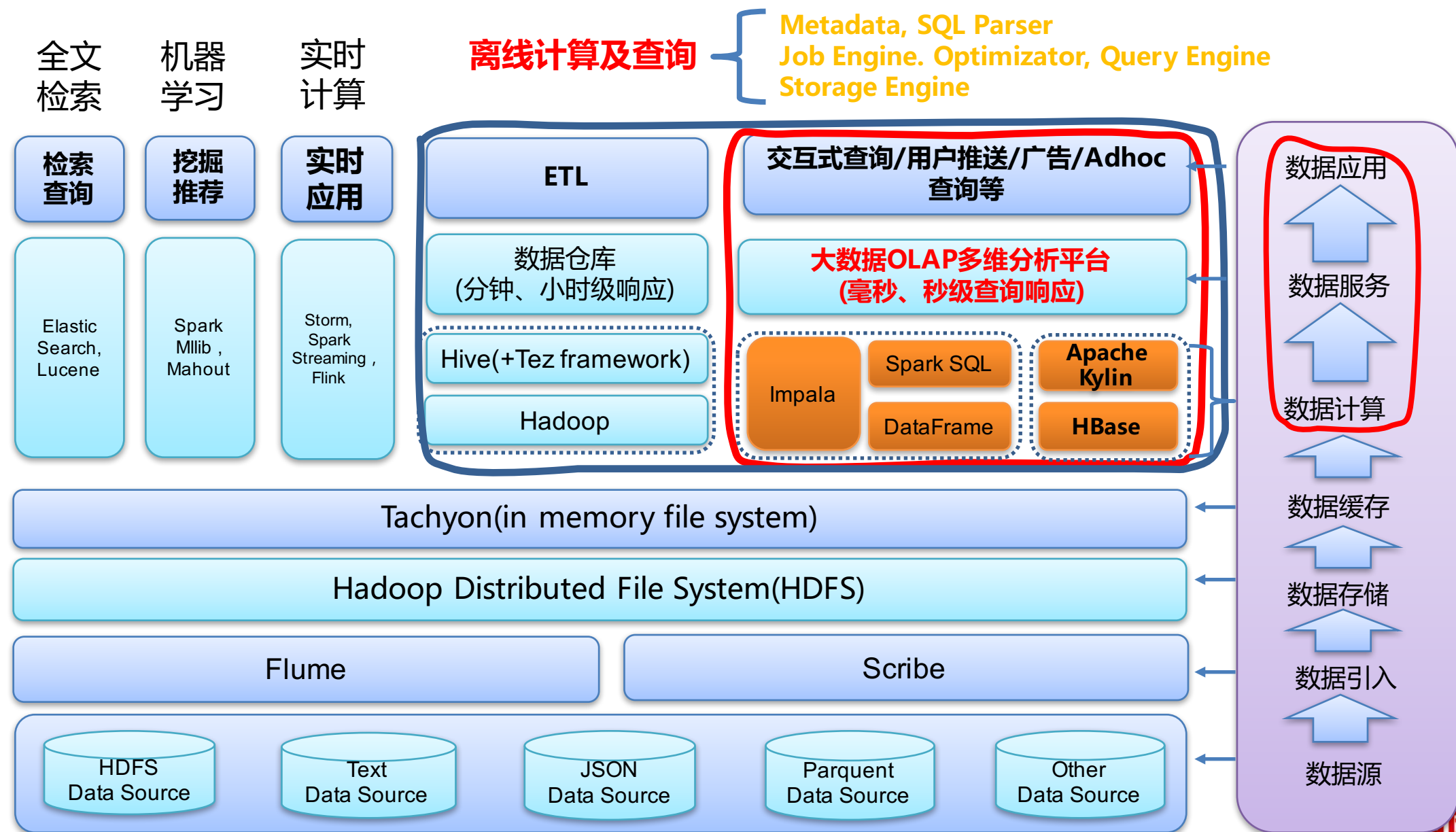
讨论

大数据多维分析平台建设背景



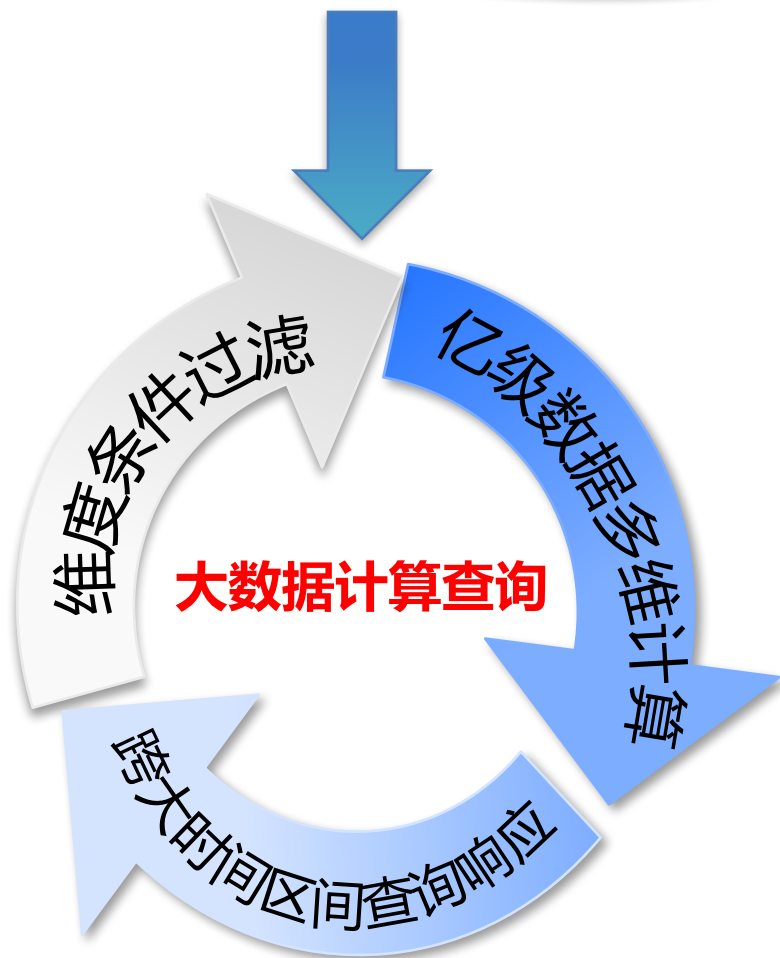
大数据OLAP多维分析平台建设

大数据查询框架选型及调研 – 相关生态



大数据查询框架选型及调研 – 技术选型

MySQL: 维度少+数据量级小+查询响应慢



大数据业界离线计算查询框架分类

1.0

基于**Runtime Framework**的架构

(如Hive, Pig)(分钟、小时级)

基于**内存计算的MPP**架构

(如Impala, Presto, Spark SQL, Palo) (秒级、分钟级)

2.0

基于**预计算和索引**的架构

(如Apache Kylin, Pinot)(毫秒、秒级)

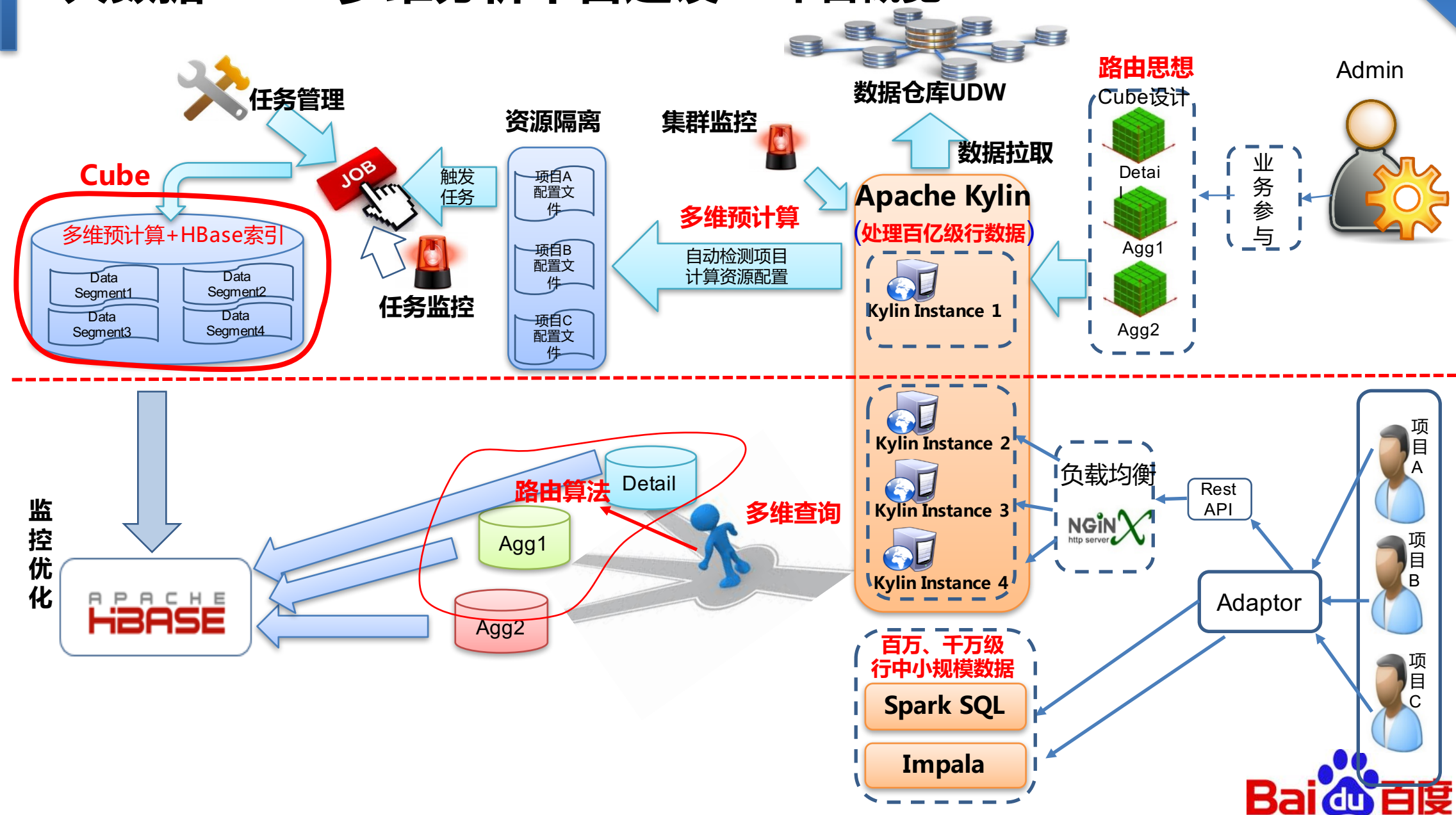
3.0

大数据OLAP多维分析平台建设

大数据OLAP多维分析平台建设 – 系统位置



大数据OLAP多维分析平台建设 - 平台概览



大数据OLAP多维分析平台建设 – 平台服务

	旧系统		新系统
系统	基于MySQL等引擎	基于Memory-Based引擎	基于Pre-calculated引擎
支持数据量级	x < 百万行级	百万行级 < x < 亿行级	千万行级 < x < 百亿行级
支持维度数量	小于5个	任意	<=15个(建议)
支持指标数量	小于5个	任意	任意
查询响应	毫秒到秒	秒到分钟	毫秒到秒
系统特点	RDBMS	内存	空间换时间

我们的深度开发

我们的深度开发 – 开发工作

平台部署

Hadoop
Hive
HBase
Kylin

01

任务调度



存储原则
计算任务
更新任务
合并任务

03

监控优化



Hadoop优化
Hive优化
HBase优化
Zookeeper优化

数据拉取依赖监控
Cube任务监控
Hadoop, HBase监控
Apache Kylin监控

05

异构数据源
依赖检测
超时预警
落地加工



数据拉取

02

项目粒度
资源隔离

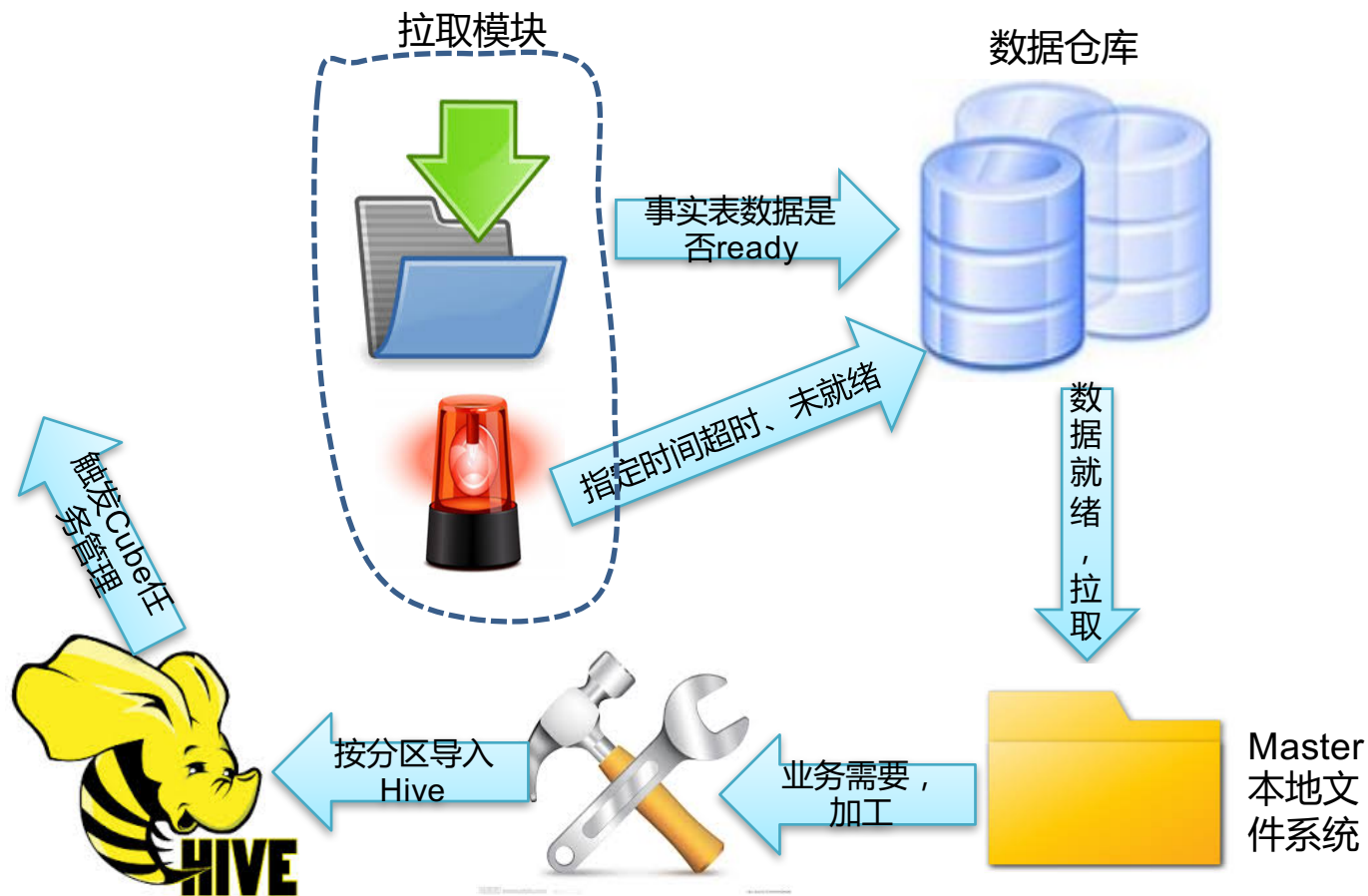


资源隔离

04

原生态暂不支持资源隔离

我们的深度开发 – 数据拉取



数据源： HDFS, MySQL等

依赖检测： 数据就绪依赖检测

超时预警： 数据在指定时间内未就绪，短信或邮件报警

落地加工： 根据业务需要，加工落地的数据，再导入Hive并计算。

我们的深度开发 – 任务调度

数据存储原则:

当月数据保持天为单位的分区，
历史数据保持月为单位的分区

计算(build)

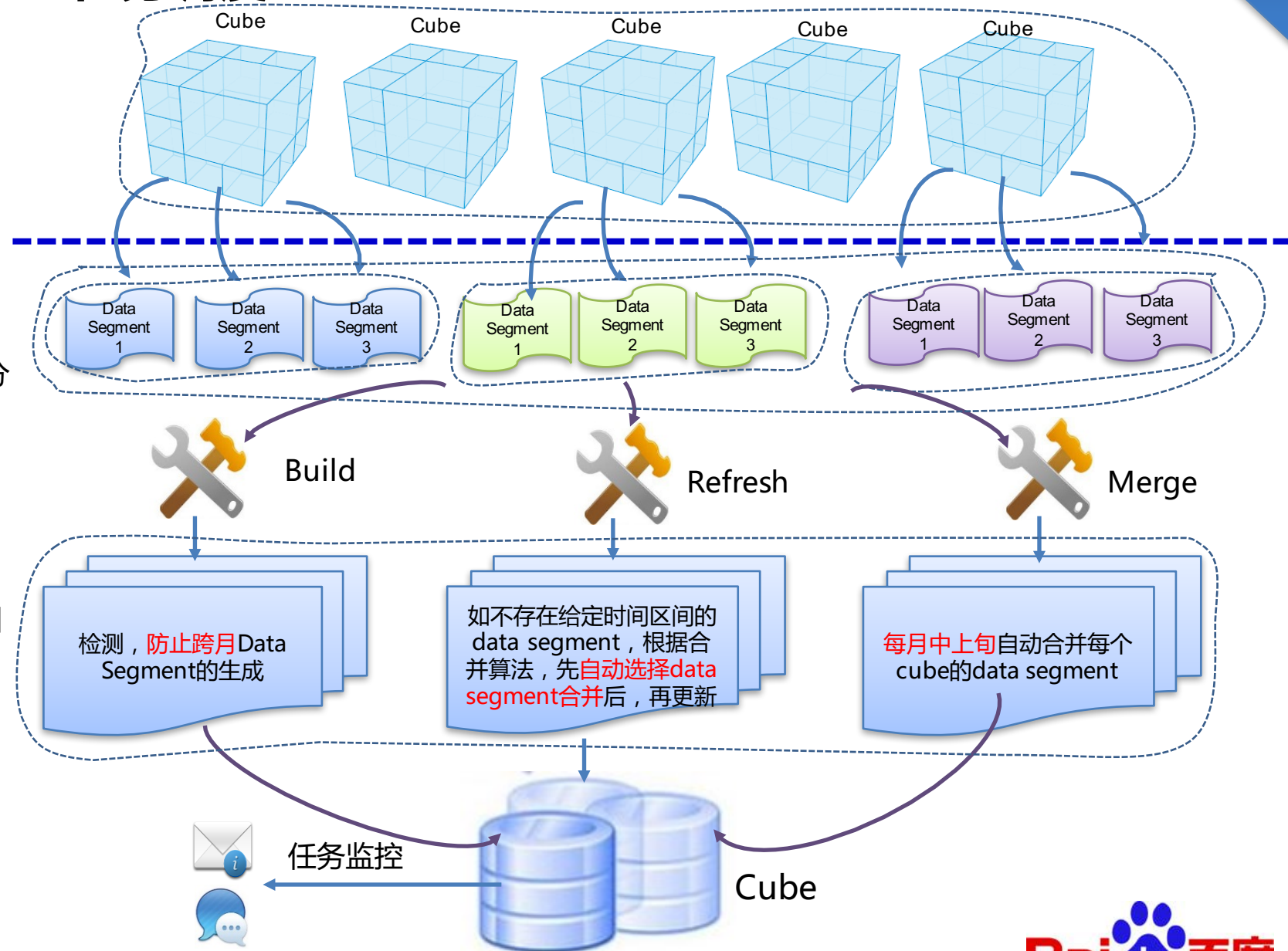
自动划分长时间线段成自然月。
源数据延迟产出可能造成数据分区跨月的问题

更新(refresh)

当月数据回溯及历史数据优化回溯策略

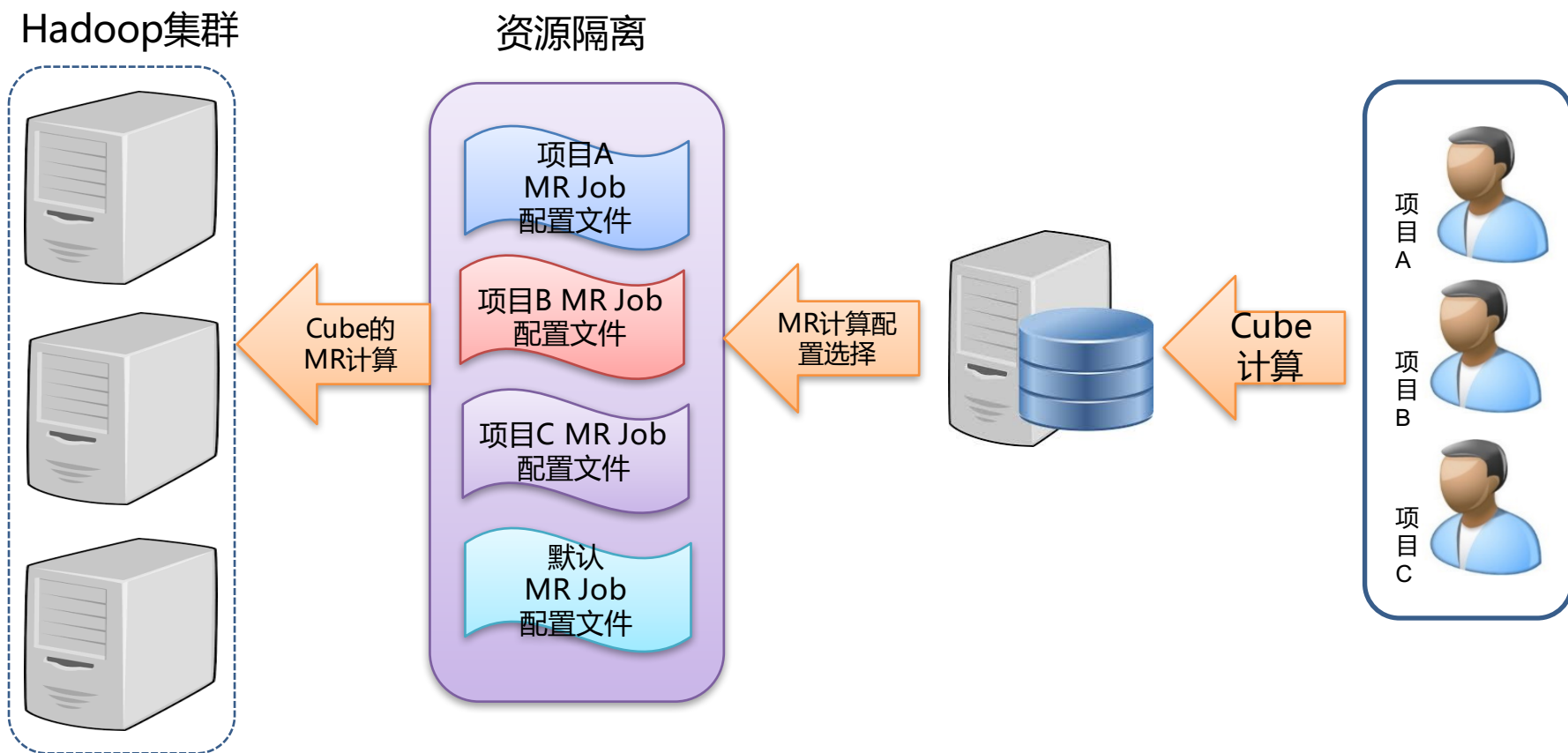
合并(merge)

每月中上旬自动合并上月所有
cube的分区

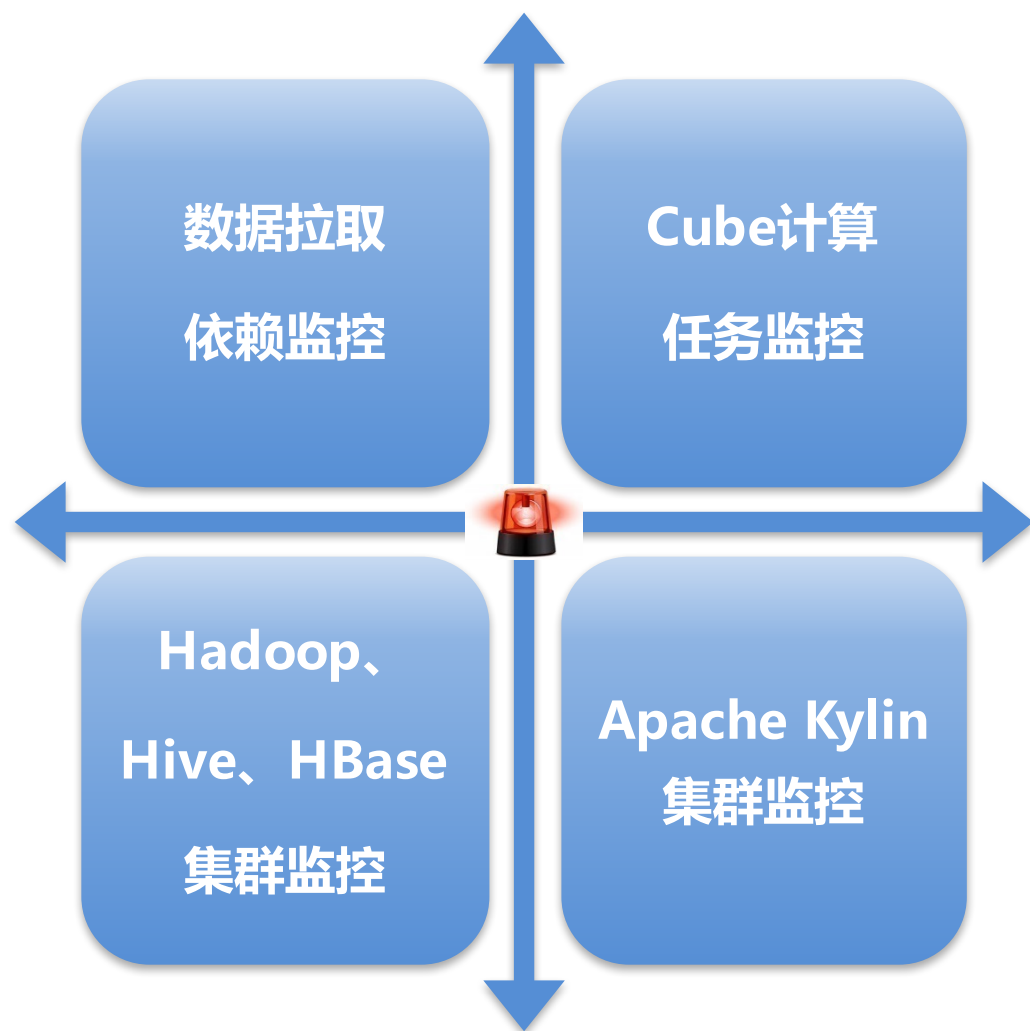


我们的深度开发 – 资源隔离

原生态暂不支持资源隔离，修改Apache Kylin代码支持项目级资源隔离



我们的深度开发 – 监控优化



性能优化



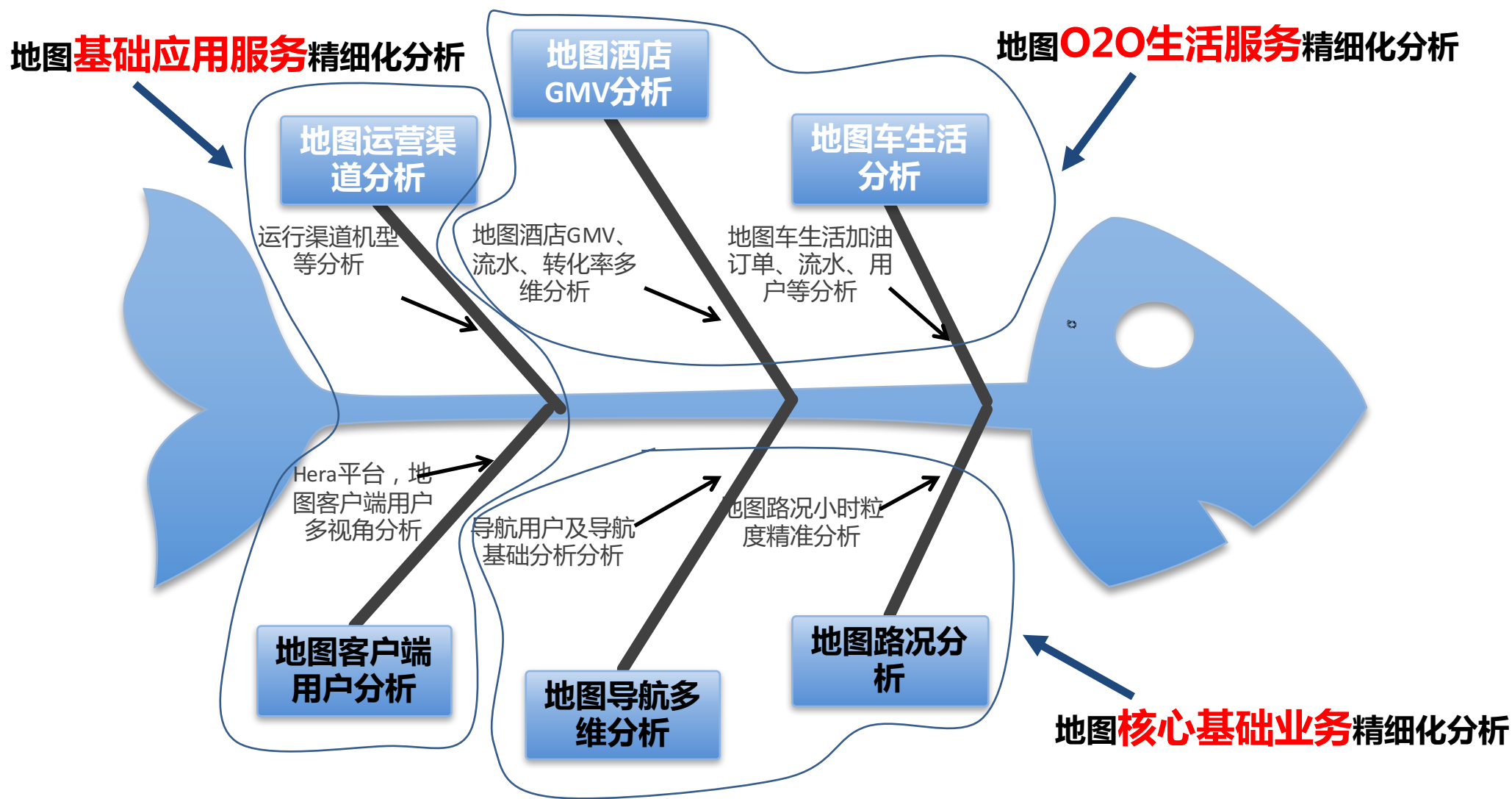
Hadoop: 要Map和Reduce的内存设置优化

HBase: Full GC优化, 关闭MajorCompaction及压缩阈值调整, 增大PermSize及ZK通信超时时间, 等, 有效减少HBaseRegionServer随机Down机问题

Zookeeper: 增大ZK session超时时间等

我们的项目实践

我们的项目实践 - 总体概览



共计**50+亿**行源数据，最大**单表20+亿**行记录

我们的项目实践 – 通用技术方案产出

1. Cube的查询性能主要因素之一即为Cube维度部分模型的设计，可根据同一维度的Hierarchy层次关系及维度的元素的秩进行设计

**CUBE
设计**

任务管理

2. Cube数据以segment为存储单位，主要设计计算、更新、分区合并等操作，目前已开发基于时间线端的自动化管理逻辑。

**不可累加
指标**

3. 对于某一维度下各元素值不互斥的情况下，向高纬度汇总累加出现数据膨胀错误，利用agg+detail cube结合路由算法实现

**留存
级联更新**

4. 留存涉及对历史多天数据的级联更新，影响数据规模大，性能降低，利用数据旋转存储，前端辅助解决，仅更新1天数据即可

**动态
JOIN
解决**

5. 数据仓库端在biglog平台做JOIN, 多维分析系统前端利用Map实现JOIN

实际项目过程中归纳产出通用技术解决方案

Q&A
谢谢